

Rec'd PCT/PTO 28 APR 2004

10/510838

## RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.

Application Serial Number: 10/501,838A

Source: PCT

Date Processed by STIC: 4-28-05

***ENTERED***



PCT

## RAW SEQUENCE LISTING

DATE: 04/28/2005

PATENT APPLICATION: US/10/501,838A

TIME: 16:18:00

Input Set : A:\24348-501NATL.ST25.txt

Output Set: N:\CRF4\04272005\J501838A.raw

3 <110> APPLICANT: Ben-Sasson, Shmuel A.  
 4 Cohen, Einat  
 6 <120> TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration  
 Across a  
 7 Biological Barrier  
 9 <130> FILE REFERENCE: 24348-501 NATL  
 11 <140> CURRENT APPLICATION NUMBER: US 10/501,838A  
 C--> 12 <141> CURRENT FILING DATE: 2004-07-19  
 14 <150> PRIOR APPLICATION NUMBER: PCT/IB03/00968  
 15 <151> PRIOR FILING DATE: 2003-02-07  
 17 <150> PRIOR APPLICATION NUMBER: US 60/355,396  
 18 <151> PRIOR FILING DATE: 2002-02-07  
 20 <160> NUMBER OF SEQ ID NOS: 72  
 22 <170> SOFTWARE: PatentIn version 3.2  
 24 <210> SEQ ID NO: 1  
 25 <211> LENGTH: 23  
 26 <212> TYPE: PRT  
 27 <213> ORGANISM: Haemophilus influenzae  
 29 <400> SEQUENCE: 1  
 31 Asn Tyr His Asp Ile Val Leu Ala Leu Ala Gly Val Cys Gln Ser Ala  
 32 1 5 10 15  
 35 Lys Leu Val His Gln Leu Ala  
 36 20  
 39 <210> SEQ ID NO: 2  
 40 <211> LENGTH: 23  
 41 <212> TYPE: PRT  
 42 <213> ORGANISM: Pasteurella multocida  
 44 <400> SEQUENCE: 2  
 46 Asn Tyr Tyr Asp Ile Thr Leu Ala Leu Ala Gly Val Cys Gln Ala Ala  
 47 1 5 10 15  
 50 Lys Leu Val Gln Gln Phe Ala  
 51 20  
 54 <210> SEQ ID NO: 3  
 55 <211> LENGTH: 23  
 56 <212> TYPE: PRT  
 57 <213> ORGANISM: Escherichia coli  
 59 <400> SEQUENCE: 3  
 61 Asn Tyr Tyr Asp Ile Thr Leu Ala Leu Ala Gly Ile Cys Gln Ser Ala  
 62 1 5 10 15  
 65 Arg Leu Val Gln Gln Leu Ala  
 66 20  
 69 <210> SEQ ID NO: 4  
 70 <211> LENGTH: 23  
 71 <212> TYPE: PRT

(pg. 6)

## RAW SEQUENCE LISTING

DATE: 04/28/2005

PATENT APPLICATION: US/10/501,838A

TIME: 16:18:00

Input Set : A:\24348-501NATL.ST25.txt

Output Set: N:\CRF4\04272005\J501838A.raw

72 <213> ORGANISM: *Vibrio cholerae*

74 &lt;400&gt; SEQUENCE: 4

76 Ala Ile Tyr Asp Arg Thr Ile Ala Phe Ala Gly Ile Cys Gln Ala Val

77 1 5 10 15

80 Ala Leu Val Gln Gln Val Ala

81 20

84 &lt;210&gt; SEQ ID NO: 5

85 &lt;211&gt; LENGTH: 23

86 &lt;212&gt; TYPE: PRT

87 <213> ORGANISM: *Buchnera aphidicola*

89 &lt;400&gt; SEQUENCE: 5

91 Lys Ile His Leu Ile Thr Leu Ser Leu Ala Gly Ile Cys Gln Ser Ala

92 1 5 10 15

95 His Leu Val Gln Gln Leu Ala

96 20

99 &lt;210&gt; SEQ ID NO: 6

100 &lt;211&gt; LENGTH: 23

101 &lt;212&gt; TYPE: PRT

102 <213> ORGANISM: *Pseudomonas aeruginosa*

104 &lt;400&gt; SEQUENCE: 6

106 Asp Pro Arg Gln Gln Leu Ile Ala Leu Gly Ala Val Phe Glu Ser Ala

107 1 5 10 15

110 Ala Leu Val Asp Lys Leu Ala

111 20

114 &lt;210&gt; SEQ ID NO: 7

115 &lt;211&gt; LENGTH: 23

116 &lt;212&gt; TYPE: PRT

117 <213> ORGANISM: *Xylella fastidiosa*

119 &lt;400&gt; SEQUENCE: 7

121 Leu Ile Asp Asn Arg Val Leu Ala Leu Ala Gly Val Val Gln Ala Leu

122 1 5 10 15

125 Gln Gln Val Arg Gln Ile Ala

126 20

129 &lt;210&gt; SEQ ID NO: 8

130 &lt;211&gt; LENGTH: 23

131 &lt;212&gt; TYPE: PRT

132 <213> ORGANISM: *Rhizobium loti*

134 &lt;400&gt; SEQUENCE: 8

136 Asn Leu Pro Pro Ile Val Leu Ala Val Ile Gly Ile Cys Ala Ala Val

137 1 5 10 15

140 Phe Leu Leu Gln Tyr Val

141 20

144 &lt;210&gt; SEQ ID NO: 9

145 &lt;211&gt; LENGTH: 23

146 &lt;212&gt; TYPE: PRT

147 <213> ORGANISM: *Homo sapiens*

149 &lt;400&gt; SEQUENCE: 9

151 Asn Tyr Phe Ile Val Asn Leu Ala Leu Ala Asp Leu Cys Met Ala Ala

152 1 5 10 15

## RAW SEQUENCE LISTING

DATE: 04/28/2005

PATENT APPLICATION: US/10/501,838A

TIME: 16:18:00

Input Set : A:\24348-501NATL.ST25.txt

Output Set: N:\CRF4\04272005\J501838A.raw

```

155 Phe Asn Ala Ala Phe Asn Phe
156      20
159 <210> SEQ ID NO: 10
160 <211> LENGTH: 23
161 <212> TYPE: PRT
162 <213> ORGANISM: Chlamydia pneumoniae
164 <400> SEQUENCE: 10
166 Thr Ala Phe Asp Phe Asn Lys Met Leu Asp Gly Val Cys Thr Tyr Val
167 1      5      10      15
170 Lys Gly Val Gln Gln Tyr Leu
171      20
174 <210> SEQ ID NO: 11
175 <211> LENGTH: 23
176 <212> TYPE: PRT
177 <213> ORGANISM: Rhizobium loti
179 <400> SEQUENCE: 11
181 Arg Ala Ile Leu Ile Pro Leu Ala Leu Ala Gly Leu Cys Gln Val Ala
182 1      5      10      15
185 Arg Ala Gly Asp Ile Ser Ser
186      20
189 <210> SEQ ID NO: 12
190 <211> LENGTH: 25
191 <212> TYPE: PRT
192 <213> ORGANISM: Bacillus subtilis
194 <400> SEQUENCE: 12
196 Met Arg Asn Leu Thr Lys Thr Ser Leu Leu Leu Ala Gly Leu Cys Thr
197 1      5      10      15
200 Ala Ala Gln Met Val Phe Val Thr His
201      20      25
204 <210> SEQ ID NO: 13
205 <211> LENGTH: 25
206 <212> TYPE: PRT
207 <213> ORGANISM: Kingella denitrificans
209 <400> SEQUENCE: 13
211 Ile Glu Leu Met Ile Val Ile Ala Ile Ile Gly Ile Leu Ala Ala Ile
212 1      5      10      15
215 Ala Leu Pro Ala Tyr Gln Glu Tyr Val
216      20      25
219 <210> SEQ ID NO: 14
220 <211> LENGTH: 25
221 <212> TYPE: PRT
222 <213> ORGANISM: Eikenella corrodens
224 <400> SEQUENCE: 14
226 Ile Glu Leu Met Ile Val Ile Ala Ile Ile Gly Ile Leu Ala Ala Ile
227 1      5      10      15
230 Ala Leu Pro Ala Tyr Gln Asp Tyr Val
231      20      25
234 <210> SEQ ID NO: 15
235 <211> LENGTH: 16

```

## RAW SEQUENCE LISTING

DATE: 04/28/2005

PATENT APPLICATION: US/10/501,838A

TIME: 16:18:00

Input Set : A:\24348-501NATL.ST25.txt

Output Set: N:\CRF4\04272005\J501838A.raw

```

236 <212> TYPE: PRT
237 <213> ORGANISM: Zonula occludens toxin
239 <400> SEQUENCE: 15
241 Ala Ser Phe Gly Phe Cys Ile Gly Arg Leu Cys Val Gln Asp Gly Phe
242 1          5          10          15
245 <210> SEQ ID NO: 16
246 <211> LENGTH: 4
247 <212> TYPE: PRT
248 <213> ORGANISM: Artificial sequence
250 <220> FEATURE:
251 <223> OTHER INFORMATION: Synthetic: Cleavable linker peptide
253 <400> SEQUENCE: 16
255 Ile Glu Gly Arg
256 1
259 <210> SEQ ID NO: 17
260 <211> LENGTH: 6
261 <212> TYPE: PRT
262 <213> ORGANISM: Artificial sequence
264 <220> FEATURE:
265 <223> OTHER INFORMATION: Synthetic: Cleavable linker peptide
267 <400> SEQUENCE: 17
269 Gly Gly Lys Gly Gly Lys
270 1          5
273 <210> SEQ ID NO: 18
274 <211> LENGTH: 29
275 <212> TYPE: PRT
276 <213> ORGANISM: Artificial sequence
278 <220> FEATURE:
279 <223> OTHER INFORMATION: Synthetic: penetrating peptide
282 <220> FEATURE:
283 <221> NAME/KEY: MISC_FEATURE
284 <222> LOCATION: (26)..(29)
285 <223> OTHER INFORMATION: cleavable linker peptide
287 <220> FEATURE:
288 <221> NAME/KEY: MISC_FEATURE
289 <222> LOCATION: (26)..(29)
290 <223> OTHER INFORMATION: wherein recombinant human insulin is coupled to the
penetrating
291 peptide via the cleavable linker peptide
293 <400> SEQUENCE: 18
295 Asn Tyr Tyr Asp Ile Thr Leu Ala Leu Ala Gly Ile Cys Gln Ser Ala
296 1          5          10          15
299 Arg Leu Val Gln Gln Leu Ala Gly Gly Ile Glu Gly Arg
300          20          25
303 <210> SEQ ID NO: 19
304 <211> LENGTH: 25
305 <212> TYPE: PRT
306 <213> ORGANISM: Artificial sequence
308 <220> FEATURE:
309 <223> OTHER INFORMATION: Synthetic: penetrating peptide

```

## RAW SEQUENCE LISTING

DATE: 04/28/2005

PATENT APPLICATION: US/10/501,838A

TIME: 16:18:00

Input Set : A:\24348-501NATL.ST25.txt

Output Set: N:\CRF4\04272005\J501838A.raw

```

312 <220> FEATURE:
313 <221> NAME/KEY: MISC_FEATURE
314 <222> LOCATION: (25)..(25)
315 <223> OTHER INFORMATION: wherein recombinant human insulin is coupled to the
penetrating
316     peptide via the glycine residue
318 <400> SEQUENCE: 19
320 Asn Tyr Tyr Asp Ile Thr Leu Ala Leu Ala Gly Ile Cys Gln Ser Ala
321 1      5      10      15
324 Arg Leu Val Gln Gln Leu Ala Gly Gly
325     20      25
328 <210> SEQ ID NO: 20
329 <211> LENGTH: 30
330 <212> TYPE: PRT
331 <213> ORGANISM: Artificial sequence
333 <220> FEATURE:
334 <223> OTHER INFORMATION: Synthetic: penetrating peptide
337 <220> FEATURE:
338 <221> NAME/KEY: MISC_FEATURE
339 <222> LOCATION: (26)..(29)
340 <223> OTHER INFORMATION: cleavable linker peptide
342 <220> FEATURE:
343 <221> NAME/KEY: MISC_FEATURE
344 <222> LOCATION: (30)..(30)
345 <223> OTHER INFORMATION: wherein heparin is coupled to the penetrating peptide via
the
346     free amino group of the lysine residue
348 <400> SEQUENCE: 20
350 Asn Tyr Tyr Asp Ile Thr Leu Ala Leu Ala Gly Ile Cys Gln Ser Ala
351 1      5      10      15
354 Arg Leu Val Gln Gln Leu Ala Gly Gly Ile Glu Gly Arg Lys
355     20      25      30
358 <210> SEQ ID NO: 21
359 <211> LENGTH: 26
360 <212> TYPE: PRT
361 <213> ORGANISM: Artificial sequence
363 <220> FEATURE:
364 <223> OTHER INFORMATION: Synthetic: penetrating peptide
367 <220> FEATURE:
368 <221> NAME/KEY: MISC_FEATURE
369 <222> LOCATION: (26)..(26)
370 <223> OTHER INFORMATION: wherein heparin is coupled to the penetrating peptide via
the
371     free amino group of the lysine residue
373 <400> SEQUENCE: 21
375 Asn Tyr Tyr Asp Ile Thr Leu Ala Leu Ala Gly Ile Cys Gln Ser Ala
376 1      5      10      15
379 Arg Leu Val Gln Gln Leu Ala Gly Gly Lys
380     20      25
383 <210> SEQ ID NO: 22
384 <211> LENGTH: 30
385 <212> TYPE: PRT

```

RAW SEQUENCE LISTING ERROR SUMMARY      DATE: 04/28/2005  
PATENT APPLICATION: US/10/501,838A      TIME: 16:18:01

Input Set : A:\24348-501NATL.ST25.txt  
Output Set: N:\CRF4\04272005\J501838A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:38; Xaa Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16  
Seq#:39; Xaa Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22  
Seq#:39; Xaa Pos. 23  
Seq#:40; Xaa Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22  
Seq#:40; Xaa Pos. 23  
Seq#:41; Xaa Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22  
Seq#:41; Xaa Pos. 23  
Seq#:42; Xaa Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22  
Seq#:42; Xaa Pos. 23,24,25  
Seq#:43; Xaa Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22  
Seq#:43; Xaa Pos. 23  
Seq#:44; Xaa Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22  
Seq#:44; Xaa Pos. 23  
Seq#:45; Xaa Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22  
Seq#:45; Xaa Pos. 23  
Seq#:46; Xaa Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22  
Seq#:46; Xaa Pos. 23,24  
Seq#:47; Xaa Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22  
Seq#:47; Xaa Pos. 23,24,25  
Seq#:48; Xaa Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22  
Seq#:48; Xaa Pos. 23  
Seq#:49; Xaa Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22  
Seq#:49; Xaa Pos. 23  
Seq#:50; Xaa Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22  
Seq#:50; Xaa Pos. 23,24,25,26  
Seq#:51; Xaa Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22  
Seq#:51; Xaa Pos. 23,24,25,26,27,28  
Seq#:52; Xaa Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22  
Seq#:52; Xaa Pos. 23

## VERIFICATION SUMMARY

DATE: 04/28/2005

PATENT APPLICATION: US/10/501,838A

TIME: 16:18:01

Input Set : A:\24348-501NATL.ST25.txt

Output Set: N:\CRF4\04272005\J501838A.raw

L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:934 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:38 after pos.:0  
L:1034 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:39 after pos.:0  
M:341 Repeated in SeqNo=39  
L:1118 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:40 after pos.:0  
M:341 Repeated in SeqNo=40  
L:1202 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:41 after pos.:0  
M:341 Repeated in SeqNo=41  
L:1271 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:42 after pos.:0  
M:341 Repeated in SeqNo=42  
L:1365 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:43 after pos.:0  
M:341 Repeated in SeqNo=43  
L:1429 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:44 after pos.:0  
M:341 Repeated in SeqNo=44  
L:1513 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:45 after pos.:0  
M:341 Repeated in SeqNo=45  
L:1627 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:46 after pos.:0  
M:341 Repeated in SeqNo=46  
L:1726 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:47 after pos.:0  
M:341 Repeated in SeqNo=47  
L:1820 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:48 after pos.:0  
M:341 Repeated in SeqNo=48  
L:1899 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:49 after pos.:0  
M:341 Repeated in SeqNo=49  
L:2015 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:50 after pos.:0  
M:341 Repeated in SeqNo=50  
L:2129 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:51 after pos.:0  
M:341 Repeated in SeqNo=51  
L:2208 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:52 after pos.:0  
M:341 Repeated in SeqNo=52